



B8

(21) (A1) 2,141,412  
(86) - 1993/08/02  
(43) 1994/02/17

(51) Int.Cl. <sup>6</sup> C12N 15/72; C12N 1/21

(19) (CA) **APPLICATION FOR CANADIAN PATENT** (12)

(54) Cloning and/or Sequencing Vector

(72) Bernard, Philippe - Belgium ;  
Gabant, Philippe - Belgium ;

(71) Université Libre de Bruxelles - Belgium ;

(30) (BE) 92/00696 1992/07/31

(57) 17 Claims

6,021,4/41

**Notice:** This application is as filed and may therefore contain an incomplete specification.

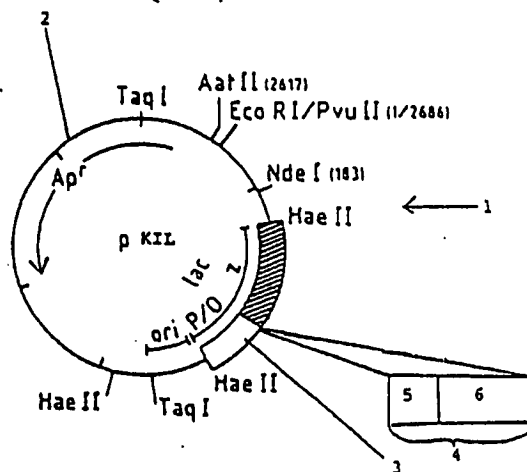


## DEMANDE INTERNATIONALE PUBLIEE EN VERTU DU TRAITE DE COOPERATION EN MATIERE DE BREVETS (PCT)

(51) Classification internationale des brevets <sup>5</sup> : <b>C12N 15/64, 15/70, 1/21 //</b> <b>(C12N 1/21, C12R 1:19)</b>	<b>A3</b>	(11) Numéro de publication internationale: <b>WO 94/03616</b> (43) Date de publication internationale: 17 février 1994 (17.02.94)
<p>(21) Numéro de la demande internationale: <b>PCT/BE93/00051</b></p> <p>(22) Date de dépôt international: <b>2 août 1993 (02.08.93)</b></p> <p>(30) Données relatives à la priorité: <b>9200696</b> <b>31 juillet 1992 (31.07.92)</b> <b>BE</b></p> <p>(71) Déposant (pour tous les Etats désignés sauf US): <b>UNIVERSITE LIBRE DE BRUXELLES (BE/BE); Avenue F.D. Roosevelt 50 CP 165, B-1050 Bruxelles (BE).</b></p> <p>(72) Inventeurs: et (75) Inventeurs/Déposants (US seulement): <b>BERNARD, Philippe (BE/BE); 52, rue Cluysenaar, B-1060 Bruxelles (BE). GABANT, Philippe (BE/BE); 9, rue Van Bemmel, B-1030 Bruxelles (BE).</b></p>	<p style="text-align: center;"><b>2141412</b></p> <p>(74) Mandataire: <b>VAN MALDEREN, Michel; Office Van Malderen, Avenue J.-S. Bach 22/43, B-1080 Bruxelles (BE).</b></p> <p>(81) Etats désignés: <b>AU, BB, BG, BR, BY, CA, CZ, FI, HU, JP, KP, KR, KZ, LK, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SK, UA, US, VN, brevet européen (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</b></p> <p>Publiée <i>Avec rapport de recherche internationale. Avant l'expiration du délai prévu pour la modification des revendications, sera republiée si de telles modifications sont reçues.</i></p> <p>(88) Date de publication du rapport de recherche internationale: <b>31 mars 1994 (31.03.94)</b></p>	

(54) Title: CLONING AND/OR SEQUENCING VECTOR

(54) Titre: VECTEUR DE CLONAGE ET/OU DE SEQUENÇAGE



## (57) Abstract

A cloning and/or sequencing vector (1) comprises, incorporated in an autonomous replication vector (2), at least one nucleotide promoter sequence (3) and at least one nucleotide sequence (4) coding for a fusion protein active as a poison. Said nucleotide sequence (4) is obtained by fusion of a coding nucleotide sequence (5), comprising several unique cloning sites and a nucleotide sequence (6) coding for a poison protein. The vector host cell of the invention is also disclosed.

## (57) Abrégé

La présente invention concerne un vecteur de clonage et/ou de séquençage (1) comprenant incorporé dans un vecteur à répllication autonome (2), au moins une séquence nucléotidique promotrice (3) et au moins une séquence nucléotidique (4) codant pour une protéine de fusion active en tant que poison; ladite séquence nucléotidique (4) étant obtenue par la fusion d'une séquence nucléotidique codante (5), comprenant plusieurs sites uniques de clonage et d'une séquence nucléotidique (6) codant pour une protéine poison. La présente invention concerne également la cellule hôte du vecteur selon l'invention.

CLONING AND/OR SEQUENCING VECTORSubject-matter of the invention

5 The invention relates to a cloning and/or sequencing vector which enables recombinant clones to be selected directly.

The invention also relates to the procaryotic cell which is transformed by this vector and to the procaryotic host cell for this vector, as well as to the use of this cloning and sequencing vector for selecting  
10 and sequencing recombinant clones.

State of the art and technological background underlying the invention

Phage (the M13 series) and plasmid (the pUC series) cloning vectors, containing numerous unique  
15 cloning sites, were constructed by Messing et al (P.N.A.S. USA, 79, pp. 3642-3646 (1977), by Norrander et al (Gene, 26, pp. 101-106 (1983) and Yanisch-Perron et al (Gene, 33 pp. 103 to 119) (1985)).

The multiple cloning sites (MCS - multiple  
20 cloning sites) of these vectors are located in the coding sequence of the LacZ gene.

Discrimination between the transformed cells which harbor a recombinant vector and the cells which harbor a non-recombinant vector is achieved using the  
25 "blue screen" technique described by Gronenborn and Messing (Methylation of single-stranded DNA in vitro introduces new restriction endonuclease cleavages sites, Nature, 272, pp. 375-377 (1978)).

However, this "blue screen" technique suffers  
30 from the disadvantage of using a screening procedure (discrimination) rather than a procedure for selecting the clones.

Discrimination by screening is based on identifying a clone within a population of clones on the  
35 basis of a characteristic (color) which differentiates it. Selection has no need of this characteristic, since it is only recombinant clones which are isolated by this method.

The screening procedure is based on the color of the recombinant clones (white color) and of the non-recombinant clones (blue color). This color is based on inactivation of the marker beta-galactosidase, preventing cleavage of X-gal (5-bromo-4-chloro-3-indolyl  $\beta$ -galactoside). The cell colonies harboring a non-recombinant vector produce a functional beta-galactosidase and, by hydrolyzing the X-gal substrate, produce a blue coloration. In general, the insertion of a DNA fragment into the  $\beta$ -galactosidase gene prevents cleavage of the X-gal. For this reason, the cells harboring a recombinant vector have a white color.

Moreover, this complex procedure requires the use of the substrate X-gal which is a product which is very expensive, unstable and awkward to use.

On the other hand, various cloning vectors permitting direct selection (positive selection) of recombinant strains have been described in the scientific literature.

Pierce et al (Proc. Natl. Acad. Sci., vol 89. No. 6, 1992, pp. 2056-2060) describe a vector which comprises the lethal gene sac B from *Bacillus amyloliquefaciens*, integrated into a plasmid derived from the bacteriophage P1 and under the control of a specific *E. coli* promoter.

The promoter of this vector includes a region having several specific cloning sites (cleavage site for a restriction enzyme).

Since the gene sac B encodes levan sucrase, which catalyzes the hydrolysis of sucrose into products which are toxic for *E. coli*, direct selection of the mutants which incorporate a recombinant plasmid is effected on a culture medium containing sucrose. Since the levan sucrase is toxic, even in the absence of sucrose, it is essential, consequently, to repress its synthesis if one wishes to obtain a large number of plasmid copies in the bacterial cytoplasm.

How ver, it is difficult, if not impossible, to repress the cytotoxic gene completely, particularly if a

large number of copies of the vector are required.

Therefore, the impossibility of repressing the cytotoxic gene leads, in phases of producing the plasmid, to the death of the cell and, as a consequence, to selective pressure towards mutated strains (characterized by an inactive lethal gene).

In this case, in order to ensure that the enzyme encoded by the sac B gene does not kill the host cell, it is necessary to incorporate a CI repressor, which regulates the expression of this gene, into the cloning vector.

Furthermore, since sucrose is often incorporated into bacterial culture media, it will be essential to prepare media which are totally free of sucrose in order to carry out these manipulations.

Henrich et al (Gene, vol 42, No. 3, 1986, pp. 345-349) describe a vector which includes the E gene from the bacteriophage  $\phi$ X174, the said E gene being incorporated into the plasmid pUH84 under the control of the Lac promoter.

In this case, the E gene includes six unique restriction sites (located over the whole of the E gene sequence) and encodes gpE, which causes lysis of the E.coli cell. In this case, positive selection is effected when a foreign recombinant gene has been inserted into one of the restriction sites.

However, this insertion of a foreign gene into a restriction site located in the sequence of the E gene, encoding gpE, makes it more difficult to sequence the foreign gene and/or amplify it by PCR since, in this case, portions of useless sequences belonging to the E gene encoding gpE are also sequenced, amplified and characterized.

Kuhn et al (Gene, vol 42, No. 3, 1986, pp. 253-263) describe a vector which includes a large gene encoding a restriction enzyme which kills by cleaving the genome of the bacterium, the said gene being incorporated into the plasmid pKG2 under the control of the LacUV5 promoter.

The cloning vectors of the state of the art suffer from the disadvantage of having to be maintained in a host strain which includes the LacI<sup>r</sup> repressor in episomal form, or the CI repressor, in order to inactivate the promoter and prevent expression of the killer gene, thus leading to the death of the host strain.

In addition, if it is desired to use this strain to produce a large number of copies of the cloning vectors, the repressor will not be adequate for preventing either a selective pressure which modifies the cytotoxic activity of the vector or a "genetic leakage", that is to say expression of certain copies of the vector and death of the host strain.

Consequently, none of the documents of the state of the art describes a cloning vector which can incorporate large nucleotide fragments, which is easy to manipulate and which can be produced by a microorganism on an industrial scale; that is to say, which can be produced in a large number of copies by a microorganism without bringing about the death of the latter.

#### Objects of the invention

The present invention aims to supply a novel cloning and/or sequencing vector, and also its host strain, which are simple and relatively inexpensive to construct and produce, and which enable recombinant clones to be selected directly, without suffering from the disadvantages of the abovementioned state of the art.

A particular object of the present invention is to obtain a vector which permits specific and certain selection of the recombinant clones.

Another object of the present invention is directed towards obtaining a vector which permits the sequencing, amplification and/or characterization, using the same primer, of any foreign DNA fragment (whatever its size) in the recombinant clones.

An additional object of the present invention is directed towards obtaining a vector which also permits simple extraction of this foreign DNA fragment from the recombinant clone.

A final object of the present invention is directed towards obtaining a host strain for the said vector which allows a large number of copies of the said vector to be produced without bringing about selection pressure which modifies the cytotoxic activity of the  
5 said vector or causing the death of the host strain.

Characteristic elements of the invention

The invention relates to a novel cloning and/or sequencing vector which includes, incorporated into an  
10 autonomously replicating vector, at least one promoter nucleotide sequence and at least one nucleotide sequence encoding a fusion protein which is active as a poison, the said nucleotide sequence being obtained by fusing a  
15 coding nucleotide sequence which includes several unique cloning sites and a nucleotide sequence which encodes a protein poison.

Preferably, the autonomously replicating vector is a recombinant virus or a recombinant plasmid such as a pUC plasmid.

20 The promoter nucleotide sequence can comprise any promoter, which permits expression of the nucleotide sequence encoding a fusion protein which is active as a poison.

Preferably, this promoter nucleotide sequence  
25 consists of the Lac operon promoter.

According to one preferred embodiment of the invention, the unique cloning sites (MCS) of the nucleotide sequence which is fused to the nucleotide sequence which encodes the protein poison are absent from  
30 the remainder of the nucleotide sequence of the vector according to the invention.

Advantageously, the nucleotide sequence of the gene which encodes the protein poison comprises all or part of the nucleotide sequence of the wild-type gene  
35 which encodes the protein CcdB.

Preferably, the nucleotide sequence of the gene which encodes the protein poison lacks the cleavage site for the restriction enzyme SmaI.

Another aspect of the invention relates to a

procaryotic cell which is transformed with the cloning vector according to the invention.

5 The invention also relates to a procaryotic host cell for the vector according to the invention which possesses a chromosomal I<sup>q</sup> and an elevated transformation efficiency, and which possesses a mutation conferring resistance to the poison activity of the fusion protein, and/or which possesses a gene encoding a protein which is an antipoison to the fusion protein.

10 Preferably, the procaryotic host cell for the vector according to the invention possesses a mutation in the gene encoding subunit A, or in the gene encoding subunit B, of the gyrase, and conferring resistance to the fusion protein, and/or a gene which encodes the protein CcdA which is an antipoison to the fusion protein.

20 Preferentially, the procaryotic cell is an Escherichia coli cell which possesses a mutation which is responsible for replacing arginine 462 with a cysteine in the amino acid sequence of the GyrA polypeptide of the gyrase, thereby conferring resistance to the fusion protein.

Preferably, this procaryotic host cell also possesses the LacI<sup>q</sup> mutation.

25 The present invention also relates to fragments of the vector according to the invention, in particular primers for sequencing and/or amplifying (for example by PCR) the foreign nucleotide fragments inserted into the vector according to the invention.

30 Preferably, these primers consist of sequences of from 10 to 30 nucleotides which hybridize to nucleotide sequences which are situated on either side of the nucleotide sequence of the vector according to the invention which contains several unique cloning sites.

35 A final aspect of the invention relates to the use of the vector according to the invention for selecting and sequencing recombinant clones.

Brief description of the figures

- Figure 1 is a diagrammatic representation of a



cloning vector according to the present invention.

- Figure 2 represents the nucleotide sequence of the gene *ccdB*, which encodes the protein CcdB.

5       - Figures 3 and 4 represent the nucleotide sequences which encode the fusion proteins of the cloning vectors pKIL18 and pKIL19, respectively. These sequences are provided with a nucleotide sequence which contains multiple unique cleavage sites for different restriction enzymes. These cloning vectors pKIL18 and pKIL19 were  
10       obtained by in vitro recombination between the wild-type *ccdB* gene of the F plasmid and plasmids pUC18 and pUC19, respectively.

Description of a preferred embodiment of the invention

15       According to the invention, the cloning and/or sequencing vector 1 includes, incorporated into an autonomously replicating vector 2, at least one promoter nucleotide sequence 3 and at least one nucleotide sequence 4 which encodes a fusion protein which is active as a poison, the said nucleotide sequence 4 being  
20       obtained by fusing a coding nucleotide sequence 5 (or polylinker) which encompasses several (multiple) unique cloning sites (MCS), and a nucleotide sequence (6) which encodes a protein poison.

25       An autonomously replicating vector 2 is understood to mean any nucleotide construct, such as a virus or a plasmid (preferably a recombinant plasmid of the PUC series), which is capable of being introduced into a microorganism, of recombining therein and/or of replicating therein.

30       Figure 1 shows a diagrammatic representation of a cloning vector according to the present invention, which vector is constructed from a plasmid of the PUC series (pUC18 and pUC19), which is described by Norrander et al (Construction of improved M13 vectors using oligo-  
35       deoxynucleotide-directed mutagenesis, Gene, 26, pp. 101-106 (1983)) and by Yanisch-Perron et al (Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors, Gene, 33, pp. 103-119 (1985)).

A coding nucleotide sequence 5 encompassing several (multiple) unique cloning sites (MCS) is understood to mean a short coding sequence (or polylinker) which comprises several cleavage sites for restriction enzymes.

The advantage of having a polylinker in the vector according to the invention is that different cloning sites are located on a single short sequence, thereby permitting:

- 10 - rapid sequencing and amplification, using the same primers, of any DNA fragment which is inserted into this vector.
- rapid extraction of the cloned fragment, facilitated by the proximity of the restriction sites. Thus, in contrast to the state of the art, this proximity avoids sequencing, amplifying and characterizing useless fragments from other sequences of the vector according to the invention.

Nucleotide sequence 6 encoding a protein poison 20 is understood to mean any (wild-type) nucleotide structure encoding a protein which displays an activity which is naturally poisonous and specific for one or more vital functions of a host cell.

A protein poison is also characterized by the existence of an antidote or antipoison, such as the 25 proteins CcdB and CcdA, the protein Kid and its antagonist Kis, the protein PemK and its antagonist PemI, the protein Doc and its antagonist Phd, the protein Hok and its antagonist Sok, and other poison molecules which are, or are not, of plasmid origin.

30 In this case, the nucleotide sequence 6 encoding a protein poison consists of the wild-type gene CcdB, which encodes the protein CcdB (control of cell death), obtained from the ccd locus of the F plasmid (Figure 2).

35 The ccd locus of the F plasmid comprises the two wild-type genes ccdA and ccdB, also termed H and G, or letA and letD, which respectively encode proteins of 72 and 101 amino acids (Bex et al, Mini-F encoded proteins; identification of a new 10.5 kilodalton species. EMBO

J.2, 1853-1861 (1983); Miki et al, Control of cell division by sex factor F in *Escherichia coli*. I. The 42.84-43.6 F segment couples cell division of the host bacteria with replication of plasmid DNA, *J. Mol. Bio.*, 174, 605-625, (1984)).

In *Escherichia coli*, the CcdB protein of the F plasmid is a cytotoxin whose lethal activity is counteracted by the protein CcdA (Karoui et al, Ham22, a mini-F mutation which is lethal to host cell and promotes recA-dependent induction of lambdoid prophage. *EMBO J.*2, 1863-1868 (1983); Ogura and Hiraga Mini-F plasmid gene that couple host cell division to plasmid proliferation, *Proc. Natl. Acad. Sci. USA*, 80, 4784-4788 (1983); Miki et al, Control of cell division by sex factor F in *Escherichia coli*. Identification of genes for inhibitor protein and trigger protein on the 42.84-43.6F segment, *J. Mol. Biol.* 174, 627-646 (1984b)).

The molecular mechanism by which protein CcdB exerts its lethal activity has been elucidated; protein CcdB is poisonous to DNA topoisomerase II.

The type II DNA topoisomerases are essential and ubiquitous enzymes which alter the topology of the DNA by transiently introducing a double-stranded break into the DNA. During the stage of break-religation, topoisomerase II forms an intermediate complex with its DNA substrate in which the enzyme is attached covalently to the 5' end of the cleaved DNA. This transitory intermediate, in which topoisomerase II is linked covalently to the DNA, has been termed the "cleavable complex" (Wang, DNA topoisomerases. *Annu. Rev. Biochem.* 54, 665-97, 1985; Maxwell & Gellert, Mechanistic aspects of DNA topoisomerases. *Advan. Protein Chem.* 38, 69-107, 1986; Liu, DNA topoisomerase poisons as antitumor drugs, *Annu. Rev. Biochem.* 58. 351-375, 1989).

Both in eucaryotes and in procaryotes, the cleavable topoisomerase II-DNA complex is the target of powerful therapeutic agents, including the antibiotics of the "quinolone" family, which act on the gyrase (bacterial topoisomerase II), and anticancer agents

(acridines and epipodophyllotoxins), which act on the mammalian topoisomerase II. The therapeutic efficacy of the topoisomerase poisons is correlated with their ability to stabilize the cleavable complex.

5 DNA topoisomerase II is an essential enzyme in all living entities and is very conserved in the evolution of the species. The CcdB protein thus displays an activity which is potentially cytotoxic for a wide variety of procaryotic species.

10 The small size of the wild-type ccdB gene allows it to be inserted into plasmids without increasing their size excessively and consequently allows large fragments of foreign DNA to be included therein. Furthermore, given its small size, the wild-type ccdB gene of the F plasmid  
15 contains very few restriction sites; it is, therefore, simpler to preserve the uniqueness of the multiple cloning sites (MCS) which are added to it.

Unexpectedly, the inventors observed that the in-phase fusion of the nucleotide sequence 6, encoding  
20 protein CcdB, with the coding nucleotide sequence (polylinker 5), comprising several (multiple) unique cloning sites (MCS), gave a nucleotide sequence 4 which encodes a fusion protein which is active as a poison and which makes it possible, as a consequence, to produce  
25 vectors for the direct selection of recombinant plasmids (killer selection).

The plasmids which have been obtained allow doubly digested restriction fragments to be cloned in both orientations with respect to the lac promoter.  
30 Insertion of a restriction fragment into one of the unique cloning sites interrupts the genetic information of the gene fusion, leading to the synthesis of a gene fusion product which is not functional. Insertional inactivation of the gene fusion ought always to take  
35 place when a termination codon is introduced or when a change is made in the reading frame.

The cells which harbor an intact cloning vector of this nature produce a poisonous fusion prot in which is functional, and die.

Insertion of a foreign DNA fragment into one of the unique cloning sites of the gene fusion interferes with production of the poison.

5 The cells which harbor a recombinant vector will be viable while cells which harbor an intact vector will not be viable. This killer selection, by simple culture on a solid medium, makes it possible to eliminate cells which harbor a non-recombinant vector (non-viable clones) and to select recombinant clones (viable clones).

10 Example I: Construction of the plasmid pKIL19

The *ccdB* gene was amplified by PCR using, as DNA template, the plasmid pULB2208 (Bernard and Couturier, The 41 carboxy-terminal residues of the miniF plasmid CcdA protein are sufficient to antagonize the killer activity of the CcdB protein, Mol. Gen. Genet. 226, 297-15 304 (1991) as well as synthetic oligonucleotides.

The synthetic oligonucleotide sequences were selected in such a way as to create an *EcoRI* restriction site on either side of the wild-type *ccdB* gene in order to be able to reclone this gene in frame with the codons of the MCS19 multiple cloning site and to eliminate the initiation codon of the native *ccdB* gene. The DNA resulting from the PCR reaction was digested with the enzyme *EcoRI* and cloned into the *EcoRI* site of the plasmid pUC19. The resulting plasmid, in which the *EcoRI* fragment was integrated in the orientation which permitted the *ccdB* gene, provided with the additional codons corresponding to the MCS19 multiple cloning sites, to be read from the *Lac* promoter, was termed pKIL2. Plasmid pKIL2 is lethal for a wild-type bacterium (*CcdB*<sup>s</sup> sensitive). 20 25 30

pKIL2 also possesses two *SmaI* sites, one in the multiple cloning sites and the other in the central region of the *ccdB* gene. The latter was eliminated by site-directed mutagenesis. The resulting plasmid pKIL3, having a unique *SmaI* site, still has two *EcoRI* sites. The *EcoRI* site downstream of the *ccdB* gene was eliminated by filling in its cohesive ends. 35

The resulting plasmid, pKIL19 (Figur 3), thus possesses a unique *EcoRI* restriction site within sequence

5, which encompasses the multiple cloning site.

Example II : construction of the plasmid pKIL18

The *ccdB* gene was amplified by PCR using, as DNA template, plasmid pKIL19 as well as synthetic oligo-nucleotides. The sequences of the synthetic oligo-nucleotides were selected in such a way as to create a HindIII site on either side of the *ccdB* gene in order to be able to reclone this gene in frame with the codons of the MCS18 multiple cloning sites. The DNA resulting from the PCR reaction was digested by the enzyme HindIII and cloned into the HindIII site of the plasmid pUC18. The resulting plasmid, in which the HindIII fragment was integrated in the orientation which permitted the *ccdB* gene, provided with the additional codons corresponding to the MCS18 multiple cloning sites, to be read from the *Lac* promoter, was termed pKIL4. Plasmid pKIL4 is lethal for a *CcdB*<sup>s</sup>-sensitive bacterium.

The HindIII site downstream of the *ccdB* gene was eliminated by filling in its cohesive ends. The resultant plasmid, pKIL18 (Figure 4), possesses a unique HindIII restriction site as well as a unique *Sma*I site (since constructed from pKIL19).

Example III : Construction of the strains *CcdB*<sup>r</sup> and *CcdB*<sup>s</sup>

In order to be able to maintain plasmids pKIL18 and pKIL19 within a bacterium, the latter has to be resistant to the lethal effect of the fusion protein which is active as a poison. Unexpectedly, the chromosomal mutation *gyrA462* confers on the strains total resistance to the poisonous effect of the fusion protein.

Moreover, since plasmids pKIL18 and pKIL19 derive directly from plasmids pUC18 and pUC19 and express the *ccdB* genes from the *Lac* promoter, it is preferable to maintain these plasmids in a *LacI*<sup>r</sup> strain. Thus, while, in our case, continuous overexpression of these genes does not exert a selection pressure in favor of certain mutations, the *LacI*<sup>r</sup> strain allows expression from the *Lac* promoter to be reduced and conserves the bacterial machinery, thereby guaranteeing a rapid generation time (increased production of the vector by the strain).

The strain D1210 (Sadler et al Gene 8, pp. 279-300 (1980)), derived from the strain HB101 *LacI<sup>r</sup>*, *LacY<sup>r</sup>* (Maniatis et al Molecular Cloning Laboratories Man. Cold Spring Harbor Laboratory N.Y.), and characterized by a chromosomal *I<sup>r</sup>* and increased transformation efficiency, was transformed with the plasmid pCOS2.1. This plasmid, which confers resistance to kanamycin, carries the *recA* gene from *Erwinia chrysanthemi* 3665 and allows recombination in *E. coli*. A lysate of P1 phage was prepared on a *CcdB<sup>r</sup> gyrA462, zei298::Tn10* strain and used to infect the strain D1210/pCOS2.1. The transductants which were resistant to tetracycline were selected and tested for their resistance or sensitivity to the *CcdB* protein. One of the *CcdB<sup>r</sup>* transductants was then cured of plasmid pCOS2.1 and termed KIB22.

Strain KIB22 constitutes an ideal host strain for plasmids pKIL18 and pKIL19 while strain D1210 constitutes the ideal host for selecting recombinant plasmids.

Thus, strain KIB22 advantageously possesses an elevated efficiency of DNA extraction (comparable to the yield of the pUC plasmids) and, unexpectedly, resistance to the fusion protein which is encoded by pKIL18 and pKIL19.

Consequently, it is possible to use this microorganism to produce the cloning vector according to the invention on an industrial scale in numerous copies without causing the death of the said microorganism.

The selection is carried out simply by spreading the bacteria on a medium containing IPTG (Isopropyl-Beta-D-thiogalactopyranoside) as well as ampicillin.

Strain KIB22 was deposited with the Laboratorium voor Microbiologie-Bacteriënverzameling (LMG) [Microbiological Laboratory - Bacterial Collection] of the Belgian Coordinated Collections of Microorganisms (BCCM) under No. LMG P-12601.

The cloning vector pKIL19 was deposited with the Laboratorium voor Moleculaire Biologie-Plasmiden Collectie (LMBP) [Molecular Biological Laboratory - Plasmid Collection] of the Belgian Coordinated

2141412

WO 94/03616

- 14 -

PCT/BE93/00051

Collections of Microorganisms (BCCM) under the No. LMBP  
2781.

5      These depositions were made in accordance with  
the provisions of the Budapest Treaty regarding the  
International Recognition of the Deposition of Microor-  
ganisms.



2141412

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: UNIVERSITE LIBRE DE BRUXELLES
- (B) STREET: 50 av. F. Roosevelt
- (C) CITY: Brussels
- (E) COUNTRY: BELGIUM
- (F) POSTAL CODE (ZIP): 1050
- (G) TELEPHONE: 32 2 650.31.00
- (H) TELEFAX: 32 2 650.35.12

(ii) TITLE OF INVENTION: Cloning vector

(iii) NUMBER OF SEQUENCES: 3

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 306 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA (genomic)

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: ccdB gene of plasmid F

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
ATGCAGTTTA AGGTTTACAC CTATAAAGA GAGAGCCGTT ATCGTCTGTT TGTGGATGTA      60
CAGAGTGATA TTATTGACAC GCCCGGGCGA CGGATGGTGA TCCCCTGGC CAGTGCACGT      120
CTGCTGTCAG ATAAAGTCTC CCGTGAAC TTACCCGGTGG TGCATATCGG GGATGAAAGC      180
TGGCGCATGA TGACCACCGA TATGGGCAGT GTGCCGGTCT CCGTTATCGG GGAAGAAGTG      240
GCTGATCTCA GCCACCGCGA AAATGACATC AAAAACGCCA TTAACCTGAT GTTCTGGGGA      300
ATATAA                                         306
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 420 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

2141412

(D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA (genomic)

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: ccdB gene of pKIL 18

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Miki,

(B) TITLE: Mini-F encoded proteins

(C) JOURNAL: EMBO

(D) VOLUME: J.2

(F) PAGES: 1853-1861

(G) DATE: 1983

(K) RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 306

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Bernard, P.

Couturier, M.

(B) TITLE: The 41 carboxy-terminal residues of the miniF plasmid CcdA protein are sufficient to antagonize the killer activity of the CcdB protein

(C) JOURNAL: Mol. Gen. Genet.

(D) VOLUME: 226

(F) PAGES: 297-304

(G) DATE: 1991

(K) RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 306

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

ATGACCATGA TTACGAATTC GAGCTCGGTA CCCGGGGATC CTCTAGAGTC GACCTGCAGG	60
CATGCAAGCT TGTCTTTGCA GTTTAAGGTT TACACCTATA AAAGAGAGAG CCGTTATCGT	120
CTGTTTGTGG ATGTACAGAG TGATATTATT GACACGCCCC GGCGACGGAT GGTGATCCCC	180
CTGGCCAGTG CACGTCTGCT GTCAGATAAA GTCTCCCGTG AACTTTACCC GGTGGTGCAT	240
ATCGGGGATG AAAGCTGGCG CATGATGACC ACCGATATGG CCAGTGTGCC GGTCTCCGTT	300
ATCGGGGAAG AAGTGGCTGA TCTCAGCCAC CGCGAAAATG ACATCAAAAA CGCCATTAC	360
CTGATGTTCT GGGGAATATA AATGTCAGGC TCCGTTATAC ACAAGCTAGC TTGGCACTGG	420

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 416 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA (genomic)

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

2141412

(A) ORGANISM: ccdB gene of pKIL 19

(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATGACCATGA TTACGCCAAG CTGTCATGCC TGCAGGTCGA CTCTAGAGGA TCCCCGGGTA	60
CCGAGCTCGA ATTCATTGCA GTTTAAGGTT TACACCTATA AAAGAGAGAG CCGTTATCGT	120
CTGTTTGTGG ATGTACAGAG TGATATTATT GACACGCCGG GCGACGGAT GGTGATCCCC	180
CTGGCCAGTG CACGTCTGCT GTCAGATAAA GTCTCCCGTG AACTTTACCC GGTGGTGCAT	240
ATCGGGGATG AAAGCTGGCG CATGATGACC ACCGATATGG CCAGTGTGCC GGTCTCCGTT	300
ATCGGGGAAG AAGTGGCTGA TCTCAGCCAC CGCGAAAATG ACATCAAAA CGCCATTAAC	360
CTGATGTTCT GGGGAATATA AATGTCAGGC TCCGTTATAC ACGAATTAAT TCAGTG	416

CLAIMS

1. Cloning and/or sequencing vector which is characterized in that it includes, incorporated into an autonomously replicating vector (2), at least one promoter nucleotide sequence (3) and at least one nucleotide sequence (4) which encodes a fusion protein which is active as a poison, the said nucleotide sequence (4) being obtained by fusing a coding nucleotide sequence (5), which comprises several cloning sites, and a nucleotide sequence (6) which encodes a protein poison.
2. Vector according to Claim 1, characterized in that the autonomously replicating vector (2) is a recombinant virus.
3. Vector according to Claim 1, characterized in that the autonomously replicating vector (2) is a recombinant plasmid.
4. Vector according to Claim 3, characterized in that the autonomously replicating vector (2) is a recombinant pUC plasmid.
5. Vector according to any one of the preceding claims, characterized in that the promoter nucleotide sequence (3) consists of the *Lac* operon promoter.
6. Vector according to any one of the preceding claims, characterized in that the unique cloning sites of the nucleotide sequence (5), which is fused to the nucleotide sequence (6) which encodes the protein poison, are absent on the remainder of the nucleotide sequence of the cloning vector.
7. Vector according to any one of the preceding claims, characterized in that the nucleotide sequence (6) which encodes a protein poison comprises all or part of the nucleotide sequence of the wild-type gene which encodes the protein CcdB.
8. Vector according to Claim 7, characterized in that the nucleotide sequence (6) which encodes the protein poison lacks the cleavage site for the restriction enzyme *Sma*I.
9. Vector according to any one of the preceding

WO 94/03616

PCT/BE93/00051

claims, characterized in that it corresponds to deposition No. LMBP2781.

10. Procaryotic cell which is transformed with the cloning vector according to any one of the preceding claims.
11. Procaryotic host cell for the cloning vector according to any one of the preceding Claims 1 to 9, characterized in that it preferably possesses a chromosomal  $I^+$  and an increased transformation efficiency and possesses a mutation which confers resistance to the poisonous activity of the fusion protein and/or possesses a gene which encodes a protein which is an antipoison to the fusion protein.
12. Procaryotic host cell for the cloning vector according to Claim 11, characterized in that it possesses a mutation in the gene encoding subunit A or in the gene encoding subunit B of the gyrase, conferring resistance to the fusion protein, and/or possesses a gene which encodes the protein CcdA which is the antipoison to the fusion protein.
13. Escherichia coli cell according to Claim 11 or 12, characterized in that it possesses a mutation which is responsible for replacing arginine 462 with a cysteine in the amino acid sequence of the GyrA polypeptide of the gyrase, thereby conferring resistance to the fusion protein.
14. Cell according to any one of Claims 11 to 13, characterized in that it possesses the  $LacI^r$  mutation.
15. Cell according to any one of the preceding Claims 11 to 14, characterized in that it is deposited under No. LMGP-12601.
16. Fragments of the vector according to any one of Claims 1 to 9, characterized in that they consist of sequences of from 10 to 30 nucleotides which hybridize to sequences situated on either side of the nucleotide sequence (5) which comprises several unique cloning sites.
17. Use of the cloning vector according to any one of Claims 1 to 9 for selecting recombinant clones.

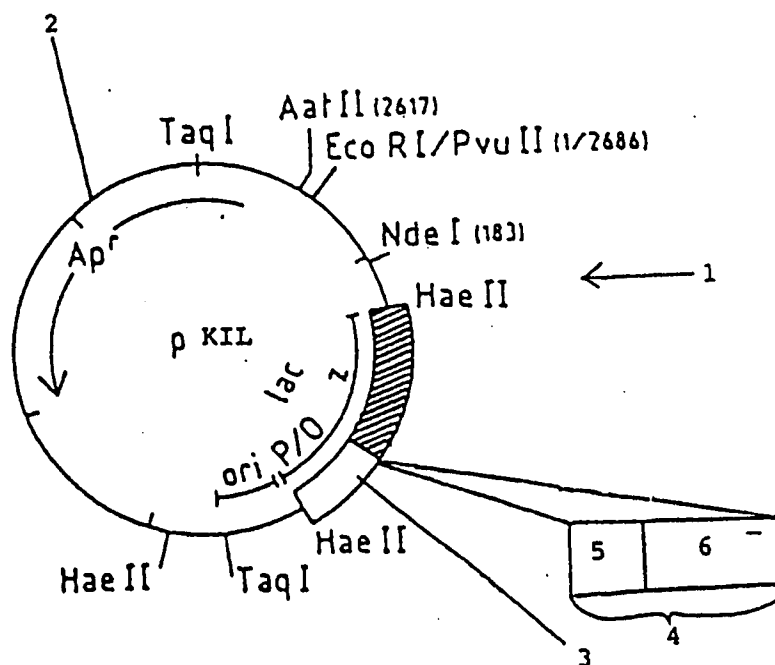


FIG. 1

2141412

WO 94/03616

PCT/BE93/00051

2/4

Figure 2

SEQ ID NO: 1

SEQUENCE TYPE: Nucleotide and its corresponding protein

SEQUENCE LENGTH: 306 base pairs

MOLECULE TYPE: Plasmid DNA

PROPERTIES: ccdB gene and the CcdB protein of the F plasmid

REFERENCE: Miki et al. 1983a

FIG.2

ATG CAG TTT AAG GTT TAC ACC TAT AAA AGA GAG AGC CGT TAT CGT CTG TTT GTG 54  
Met Gln Phe Lys Val Tyr Thr Tyr Lys Arg Glu Ser Arg Tyr Arg Leu Phe Val  
5 10 15

GAT GTA CAG AGT GAT ATT ATT GAC ACG CCC GGG CGA CGG ATG GTG ATC CCC CTG 108  
Asp Val Gln Ser Asp Ile Ile Asp Thr Pro Gly Arg Arg Met Val Ile Pro Leu  
20 25 30 35

GCC AGT GCA CGT CTG CTG TCA GAT AAA GTC TCC CGT GAA CTT TAC CCG GTG GTG 162  
Ala Ser Ala Arg Leu Leu Ser Asp Lys Val Ser Arg Glu Leu Tyr Pro Val Val  
40 45 50

CAT ATC GGG GAT GAA AGC TGG CGC ATG ATG ACC ACC GAT ATG GCC AGT GTG CCG 216  
His Ile Gly Asp Glu Ser Trp Arg Met Met Thr Thr Asp Met Ala Ser Val Pro  
55 60 65 70

GTC TCC GTT ATC GGG GAA GAA GTG GCT GAT CTC AGC CAC CGC GAA AAT GAC ATC 270  
Val Ser Val Ile Gly Glu Glu Val Ala Asp Leu Ser His Arg Glu Asn Asp Ile  
75 80 85 90

AAA AAC GCC ATT AAC CTG ATG TTC TGG GGA ATA TAA 306  
Lys Asn Ala Ile Asn Leu Met Phe Trp Gly Ile  
95

WO 94/03616

PCT/BR93/00051

3/4

Figure 3

SEQ ID NO: 2

SEQUENCE TYPE: Nucleotide and its corresponding protein

SEQUENCE LENGTH: 420 base pairs

MOLECULE TYPE: Plasmid DNA

PROPERTIES: *ccdB* gene and the CcdB protein of the plasmid

pKIL 18

FIG. 3

ATG ACCATG ATT ACG AAT TCG AGC TCG GTA CCC GGG GAT CCT CTA GAG TCG ACC 54  
Met Thr Met Ile Thr Asn Ser Ser Ser Val Pro Gly Asn Pro Leu Glu Ser Thr  
5 10 15

TGC AGG CAT GCA AGC TTG TCT TTC CAG TTT AAG GTT TAC ACC TAT AAA AGA GAG 108  
Cys Arg His Ala Lys Leu Ser Leu Gln Phe Lys Val Tyr Thr Tyr Lys Arg Glu  
20 25 30 35

AGC CGT TAT CGT CTG TTT GTG GAT GTA CAG AGT GAT ATT ATT GAC ACG CCG GGG 162  
Ser Arg Tyr Arg Leu Phe Val Asp Val Gln Ser Asp Ile Ile Asp Thr Pro Gly  
40 45 50

CGA CCG ATG GTG ATC CCC CTG GCC AGT GCA CGT CTG CTG TCA GAT AAA GTC TCC 216  
Arg Arg Met Val Ile Pro Leu Ala Ser Ala Arg Leu Leu Ser Asp Lys Val Ser  
55 60 65 70

CGT GAA CTT TAC CCG GTG GTG CAT ATC GGG GAT GAA AGC TGG CGC ATG ATG ACC 270  
Arg Glu Leu Tyr Pro Val Val His Ile Gly Asp Glu Ser Trp Arg Met Met Thr  
75 80 85 90

AAC GAT ATG GCC AGT GTG CCG GTC TCC GTT ATC GGG GAA GAA GTG GCT GAT CTC 324  
Thr Asp Met Ala Ser Val Pro Val Ser Val Ile Gly Glu Glu Val Ala Asp Leu  
95 100 105

AGC CAC CCG GAA AAT GAC ATC AAA AAC GCC ATT AAC CTG ATG TTC TGG GGA ATA 378  
Ser His Arg Glu Asn Asp Ile Lys Asn Ala Ile Asn Leu Met Phe Trp Gly Ile  
110 115 120

TAAATGTCAG GCTCC GTTAT ACACAAGCTA GCTTGCCACT GG

420



2141412

WO 94/03616

PCT/BR93/00051

4/4

FIGURE 4

SEQ ID. NO: 3

SEQUENCE TYPE: Nucleotide and its corresponding protein

SEQUENCE LENGTH: 416 base pairs

MOLECULE TYPE: plasmid DNA

PROPERTIES: *ccdB* gene and the CcdB protein of the plasmid

pKIL 19

FIG. 4

ATG ACC ATG ATT ACG CCA AGC TTG CAT GCC TGC AGG TCG ACT CTA GAG GAT CCC 54  
Met Thr Met Ile Thr Pro Ser Leu His Ala Cys Arg Ser Thr Leu Glu Asp Pro  
5 10 15

CGG GTA CCG AGC TCG AAT TCA TTG CAG TTT AAG GTT TAC ACC TAT AAA AGA GAG 108  
Arg Val Pro Ser Ser Asn Ser Leu Gln Phe Lys Val Tyr Thr Tyr Lys Arg Glu  
20 25 30 35

AGC CGT TAT CGT CTG TTT GTG GAT GTA CAG AGT GAT ATT ATT GAC ACG CCG GGG 162  
Ser Arg Tyr Arg Leu Phe Val Asp Val Gln Ser Asp Ile Ile Asp Thr Pro Gly  
40 45 50

CGA CCG ATG GTG ATC CCC CTG GGC AGT GCA CGT CTG CTG TCA GAT AAA GTC TCC 216  
Arg Arg Met Val Ile Pro Leu Ala Ser Ala Arg Leu Leu Ser Asp Lys Val Ser  
55 60 65 70

CGT GAA CTT TAC CCG GTG GTG CAT ATC GGG GAT GAA AGC TGG CGC ATG ATG ACC 270  
Arg Glu Leu Tyr Pro Val Val His Ile Gly Asp Glu Ser Trp Arg Met Met Thr  
75 80 85 90

ACC GAT ATG GCC AGT GTG CCG GTC TCC GTT ATC GGG GAA GAA GTG GCT GAT CTC 324  
Thr Asp Met Ala Ser Val Pro Val Ser Val Ile Gly Glu Glu Val Ala Asp Leu  
95 100 105

AGC CAC CGC GAA AAT GAC ATC AAA AAC GCC ATT AAC CTG ATG TTC TGG GGA ATA 378  
Ser His Arg Glu Asn Asp Ile Lys Asn Ala Ile Asn Leu Met Phe Trp Gly Ile  
110 115 120

TAAATGTCAG GCTCCGTTAT ACACGAATTA ATTCAGTG 416